A Mobile-Cloud Pedestrian Crossing Guide for the Blind

Bharat Bhargava, Pelin Angin, Lian Duan

Department of Computer Science, Purdue University
305 N University Street, West Lafayette, IN, 47907 USA

[1]bb@cs.purdue.edu
[2]pangin@cs.purdue.edu
[3]duan7@cs.purdue.edu

***Abstract*— The ability to detect the status of pedestrian signals at street intersections is a critical aspect of safe outdoor navigation for the blind and visually-impaired, especially in unfamiliar environments. Existing outdoor navigation devices for the blind prove insufficient in providing such guidance due to their reliance on limited mobile computational resources. In this work, we propose a mobile-cloud collaborative approach for context-aware outdoor navigation, where we use the computational power of resources made available by cloud computing providers for real-time image processing. The system architecture we propose also has the advantages of being extensible and having minimal infrastructural reliance, thus allowing for wide usability. We have developed an outdoor navigation application with integrated support for pedestrian crossing guidance and report experiment results, which suggest that the proposed approach is promising for real-time crossing guidance for blind pedestrians.**

***Keywords*— mobile-cloud computing; blind; assistive technology; pattern recognition; pedestrian signal**

I. Introduction

There are over 21.2 million visually impaired or blind people in the United States today [10] and many more in the world. The two biggest challenges for independent living of the blind and the visually impaired as stated in [11] are access to printed material and safe and efficient navigation. In order to navigate safely, blind people must learn how to detect obstructions, find curbs when outside and stairs when inside buildings, interpret traffic patterns, find bus stops and know their own location [11]. This implies being fully aware of the context of their living and working environment.

Crossing at urban intersections is a difficult and possibly dangerous task for the blind, hindering independent safe navigation. Assistive technology researchers have been working on this problem for years, not many of the proposed solutions being widely adopted. The currently widespread solution, also known as an accessible pedestrian signal (APS), uses a special sound/speech output to notify blind people about the status of a pedestrian signal at an intersection. However, this solution requires installation of expensive equipment at the intersections, limiting its applicability. In order to implement this solution, an extra terminal device for controlling the speaker should be added to the current traffic lights infrastructure, and this modification requires large amounts of money as well as time.

In this paper we propose an efficient and universally applicable solution for helping the blind safely cross at urban intersections. The proposed solution, based on mobile-cloud computing does not require any modification to existing pedestrian signal infrastructures, while providing guidance in real-time and being highly available.

The rest of this paper is organized as follows: Section 2 briefly mentions related work in pedestrian signal detection; Section 3 discusses the challenges involved in providing crossing guidance to blind pedestrians; Section 4 describes the mobile-cloud solution we propose for the stated problem; Section 5 gives the details of the run-time behaviour of the proposed system; Section 6 provides experiments results on the response times of the proposed system and Section 7 concludes with future work directions.

II. RELATED WORK

The problem of detecting the status of pedestrian signals at urban intersections was studied by many researchers working in the field of assistive technologies for the blind. Among some previous proposals to solve this problem are [1], [2], [3], [4], [5] and [6], which use image processing to detect the presence of different kinds of pedestrian signals in a complex scene. The major shortcoming of some of the previously proposed approaches is their reliance on a low-portability computation device for the necessary image processing and yet others are lacking the universal design principle by limiting their training data to a specific set and draining the battery of the mobile device by running the detection algorithm on the device. These approaches also require the user to take a picture/do a video recording at the intersection, although the picture taken by a blind user may not be able to capture the pedestrian signal or crossing. On the other hand, although the approach proposed by Bohonos et al [7], not relying on image processing, is promising for accurate detection, it requires installation of special hardware at pedestrian signals, which limits its use to a very small area.

III. PEDESTRIAN CROSSING GUIDANCE CHALLENGES

The inherent difficulty of providing real-time crossing guidance to blind users at intersections is the fast image processing required for locating and detecting the status of pedestrian signals in the immediate environment. As real-time image processing is demanding in terms of computational resources, mobile devices with limited resources fall short in achieving accurate and timely detection.

The problem of providing real-time feedback about the status of pedestrian signals in the immediate environment faces challenges even when a mobile-cloud collaborative approach as explained below is taken. One of the main concerns about this approach is the time it takes to send the video frames to the remote server for processing and to receive a response. The real-time nature of the problem requires response times ideally less than 1 second to provide accurate and safe guidance to the blind or visually impaired user. While the server having sufficient computation resources takes negligible time to process the received frames, network latency could create a bottleneck on the timeliness of the response to be received by the user. Continuous Internet connectivity is another problem faced by the proposed approach. Signals from wireless networks would be weak or mostly unavailable at outdoor locations, which is the main setting the application is supposed to work at. However, availability of data plans by major cell phone carrier companies today alleviates this problem. Many people are already subscribed for these data plans for a low monthly cost for continuous connectivity.

Another major challenge faced is the short battery life of the mobile device. A continuous video recording approach to the problem exhausts the battery of the mobile device too soon, causing service interruption.

The most accurate solution for detecting the status of pedestrian signals would be using a mobile device and cloud service to locate the blind user and then using the mobile device to capture the electronic signal emitted by the existing traffic lights system and send a request to the system for obtaining the time when the lights would turn green (or white based on the country). However, we found this solution had fateful difficulty on design as the electronic communication signal in existing pedestrian signal system is somehow confidential and not supported to be captured by outside applications.

IV. PROPOSED SYSTEM

The context-aware navigation system architecture we propose [12] is a two-tier architecture as seen in Figure 1. The two main components are the mobile device with integrated GPS receiver and compass, which could be any smart phone device in the market and the cloud server, which is basically the Web Services Platform that can be employed to support a variety of context-awareness functionalities. The mobile device is responsible for local navigation, local obstacle detection and avoidance, as well as interacting with the user and the cloud side. It is responsible for providing location data to the cloud, which will perform the desired location specific functionality and communicate the desired information as well as relevant context information and warnings of potential hazards in context back to the mobile device.

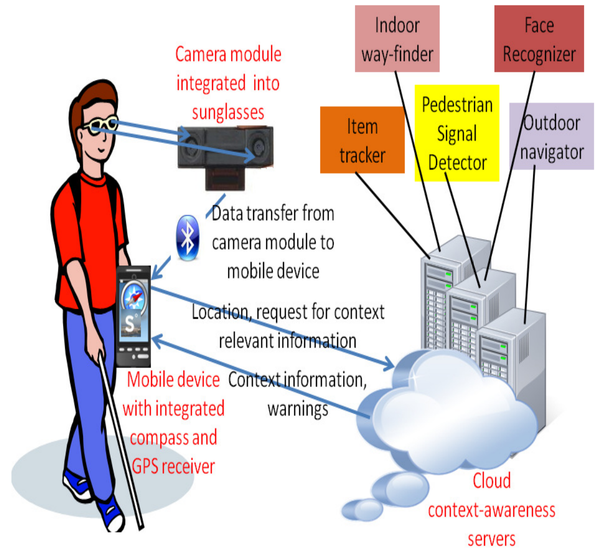

Fig. 1 Mobile-cloud architecture for context-aware navigation of the blind and visually impaired

The following subsections provide a detailed description of the major components of the proposed system:

A. Android Device with Positioning Application

Android is the mobile platform we have chosen to build context-awareness navigation functionality on, due to its open architecture, support for multi-tasking and accessibility features. Android based devices come with integrated speech recognition and text-to-speech engines, which -with some improvements- will be of help for the development of an easy-to-use interface for the blind users.

Developing our application on an Android device makes our system more universal and widely applicable. Another benefit of the Android platform is that the Android System can receive both GPS signals and 3G mobile signals, which are critical components for the location detection as well as the data transition. We extended an open source project named WalkyTalky (http://eyes-free.blogspot.com/2010/10/walking-about-with-talking-android.html) by the Eyes-free group at Google, to integrate crossing guidance into outdoor navigation. The details of this application are presented in Section 5.

B. External Camera

During the implementation of the current system, the native Android camera was used to take the pictures at street intersections. However, in future work we will be employing camera modules integrated into glasses to be worn by the user, considering the fact that the placement of the camera is of vital importance for collection of context-relevant data and an eye-level placement is the most natural. We are also considering the placement of lateral camera modules on the glasses for even better context-awareness at later stages. The best option for these modules will be the time-of-flight camera technology. Time-of-flight (TOF) range cameras are a new technology of great use in real-time three-dimensional imaging. The technology has been utilized successfully in several fields including curb and ramp detection for safe parking [13], mobile service robots for collision free manipulation of particular classes of objects [14] and obstacle detection for autonomous ground vehicles [15] and graffiti detection [16]. These cameras provide real-time depth information about pixels of a captured image, by emitting a modulated near-infrared light signal and computing the phase of the received reflected light signal [17]. A TOF camera module will not only enable us to detect nearby objects, but also those at a distance, providing greater safety especially in the case of dynamic environments with fast moving objects (such as electric cars) that the user cannot easily recognize using senses other than vision.

C. Image Processing Server on Amazon EC2 Cloud

The Elastic Compute Cloud service of Amazon Web Services (http://aws.amazon.com/ec2/) was used to host the cloud component of our system, where the server is responsible for receiving video frames from the Android mobile device, processing to detect the presence and status of pedestrian signals in the frame and sending a response as appropriate back to the mobile device over a TCP connection. We chose to use the Amazon platform in this project due to its proven robustness and ease of use.

V. SYSTEM FLOW

The different steps taken by the system to achieve crossing guidance are as follows:

1. Android Application captures GPS signal and communicates with the Google Maps Server to determine the current location of the blind user.
2. Once the application detects that the blind user is at an urban intersection and will possibly cross it, it triggers the phone camera to take a picture and send the picture to the Server running on the Amazon EC2 Cloud.
3. The server does the image processing applying the Pedestrian Signal Detection algorithm and returns the result about whether it is safe for the blind user to cross the intersection.
4. If the result is positive, the Android Application notifies the blind user to cross the intersection with speech feedback. Otherwise, it repeats step 2 until the result is positive.

Figure 2 below summarizes the system flow:

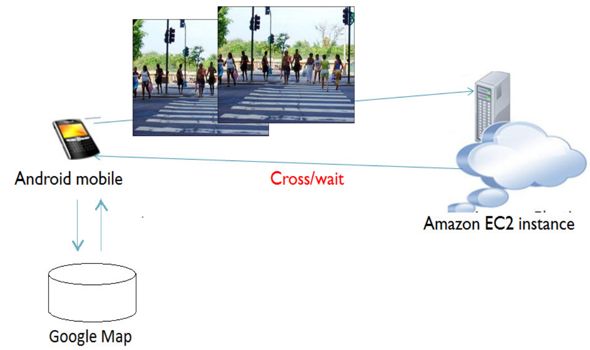

Fig. 2 Flow of actions in the proposed system

The most critical components of this system are the location detection application running on the Android Phone and the Image Processing Application running on the Amazon EC2 Cloud. Following are detailed information about these two components.

A. Positioning

For positioning, we extended the WalkyTalky navigation application, which is an open source project that can speak out the address of a nearby location passed by the user and navigates the user to the destination using GPS and Google Maps. We removed the graphical user interface of this application as it is useless for the blind people. The automatic image capturing function was added to the application as well. In order to make the detection more accurate, we enabled the Compass function in the application. The reason for using the Compass function is obtaining the heading direction of the blind. Consider the case presented in figure 3, where a user arrives at an intersection but is not heading the correct direction. With the heading direction information, the application can detect this case and warn the user to change his/her position to point exactly in the direction of the pedestrian signal so that a better picture can be taken.

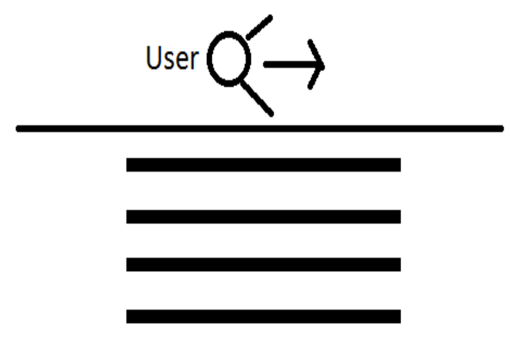

Fig. 3 Disoriented user at intersection

B. *Multi-cue Signal Detection Algorithm*

The pedestrian signal detector of the developed system uses a cascade of boosted classifiers based on the AdaBoost algorithm [8], which is popular for real-time object recognition tasks and haar-like features [9] to detect the presence and status of pedestrian signals in a picture captured by the camera of the Android mobile phone. We are currently investigating the effectiveness of a multi-cue algorithm in providing accurate guidance to the user at intersections. As seen in figure 4, the presence of contextual clues including other pedestrians crossing in the same direction, a zebra crossing and the status of traffic lights in the same direction provide additional information to make an accurate decision about whether the user should cross. With the help of Cloud Computing, we will be able to run all detection algorithms (those detecting the state of other contextual cues) in parallel to make a more informed and conservative decision at the crossing. Another important aspect we will take into consideration in development of the detection algorithm is the universal aspects of pedestrian signals. As signals in different countries and even different cities can be dramatically different from each other, it will be important to focus on the common features at the image processing stage, instead of training the detector with a dataset of signal images, which may not be comprehensive.

Fig. 4 Cues to help provide accurate crossing guidance at an intersection

(image from http://news.bbc.co.uk)

VI. EXPERIMENTAL EVALUATION

The two most important aspects of the pedestrian signal status detection problem are timeliness of response and accuracy. The real-time nature of the problem necessitates response times of less than 1 second as stated before, while high accuracy of detection should be achieved to ensure safety of the user. Experiments were performed to test the response time of the pedestrian signal detector application developed. Test data used in the experiments consists of pictures at outdoor locations of the Purdue University campus, which include scenes of different pedestrian signals. The application developed was installed on an Android mobile phone, connected to the Internet through a wireless network on campus. The sample task in the experiments involved processing five different resolution level versions of pictures. The average response times, which were determined by the time period between capturing a frame and receiving the response from the server running at Amazon Elastic Compute Cloud about the pedestrian signal status, were measured for each frame resolution level as determined by a Java platform-specific measure. A resolution level of 0.75 stands for the original frame as captured by the camera, whereas the lower resolution levels represent compressed versions of the same set of frames, where image quality falls with decreasing resolution level. The response times for different resolution levels are seen in figure 5 below. Response times for the original frames are around 660 milliseconds on average, which are acceptable levels for the real-time requirements of the problem. We also see that response time decreases further when lower-quality, compressed versions of the frames are sent to the remote server instead of the originals.

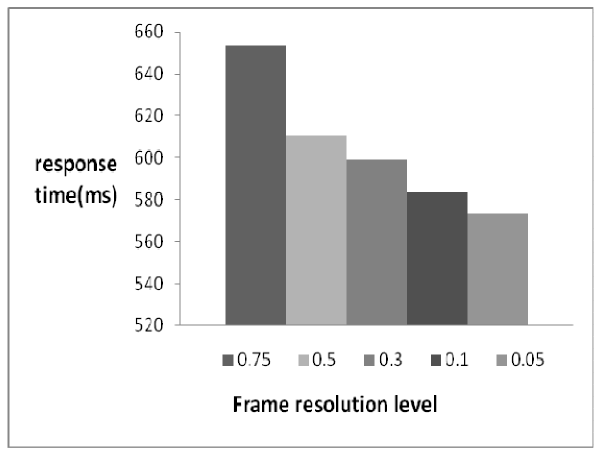

Fig. 5 Response time experiment results

VII. CONCLUSION AND FUTURE WORK

In this paper we proposed a mobile-cloud based pedestrian signal detector integrated into an outdoor navigation application for the blind and visually impaired using the Android platform for the mobile component and the Amazon EC2 platform as the cloud component. We also conducted experiments to test the appropriateness of the developed system for real-time guidance and the results are promising for wide adoption of a navigation aid based on the proposed system.

Our future work will involve development and experimentation of a system using camera modules placed on eye glasses instead of using the native camera of the mobile device for a more accurate and comfortable system. We will also investigate privacy issues arising from the sending of possibly confidential data (location and image data) to the cloud for processing and work on solutions for protecting privacy under the real-time response constraint.

REFERENCES

[1] Charette, R. and Nashashibi, F. Real Time Visual Traffic Lights Recognition Based on Spot Light Detection and Adaptive Traffic Lights Templates. In *World Congress and Exhibition on Intelligent Transport Systems and Services* (2009).

[2] Crandall, W., Brabyn, J., Bentzen, B.L., and Myers, L. Remote Infrared Signage Evaluation for Transit Stations and Intersections. *Journal of Rehabilitation Research and Development*, 36, 4 (1999), 341-355.

[3] Ivanchenko, V., Coughlan, J., and Shen, H. Detecting and locating crosswalks using a camera phone. In *Computer Vision and Pattern Recognition Workshops* (2008).

[4] Kim, Y.K., Kim, K.W., and Yang, X. Real Time Traffic Light Recognition System for Color Vision Deficiencies. In *IEEE International Conference on Mechatronics and Automation* (2007).

[5] Shioyama, T., Wu, H., Nakamura, N., and Kitawaki, S. Measurement of the length of pedestrian crossings and detection of traffic lights from image data. *MEASUREMENT SCIENCE AND TECHNOLOGY*, 13 (2002), 1450-1457.

[6] Uddin, M. and Shioyama, T. Detection of Pedestrian Crossing using Bipolarity and Projective Invariant. In *IAPR Conference on Machine Vision Applications* (2005).

[7] Bohonos, S., Lee, A., Malik, A., Thai, C., and Manduchi, R. Universal real-time navigational assistance (URNA): An urban bluetooth beacon for the blind. In *1st ACM SIGMOBILE International Workshop on Systems and Networking Support for Healthcare and Assisted Living Environments* (2007).

[8] Freund, Y. and Schapire, R.E. A Decision-Theoretic Generalization of On-line Learning and an Application to Boosting. *Journal of Computer and System Sciences*, 55 (1997), 119-139.

[9] Lienhart, R. and Maydt, J. An Extended Set of Haar-Like Features for Rapid Object Detection. In *IEEE International Conference on Image Processing* (2002).

[10] Pleis, J.R. and Lethbridge-Çejku, M. Summary health statistics for U.S. adults: National Health Interview Survey, 2006. National Center for Health Statistics, 2007.

[11] Giudice, N.A. and G.E. Legge. Blind Navigation and the Role of Technology. In A. Helal, M. Mokhtari and B. Aldulrazak, ed., The Engineering Handbook of Smart Technology for Aging, Disability and Independence. John Wiley & Sons, Hoboken, New Jersey, 2008.

[12] Angin, P., Bhargava, B., and Helal, S. A Mobile-Cloud Collaborative Traffic Lights Detector for Blind Navigation. In *Mobile Data Management*, 2010.

[13] Gallo, O., Manduchi, R., and Rafii, A. Robust Curb and Ramp Detection for Safe Parking using the Canesta TOF camera. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops* (2008).

[14] Grundmann, T., Eidenberger, R., Zoellner, R.D., Zhixing, X., Ruehl, S., Zoellner, J.M., Dillmann, R., Kuehnle, J., and Verl, A. Integration of 6D Object Localization and Obstacle Detection for Collision Free Robotic Manipulation. In *IEEE/SICE International Symposium on System Integration* (2008).

[15] Bostelman, R.V., Hong, T.H., and Madhavan, R. Towards AGV Safety and Navigation Advancement - Obstacle Detection using a TOF Range Camera. In *International Conference on Advanced Robotics* (2005).

[16] Tombari, F., Stefano, L., Mattoccia, S., and Zanetti, A. Graffiti Detection Using a Time-of-Flight Camera. In *10th International Conference on Advanced Concepts for Intelligent Vision Systems* (2008).

[17] Ringbeck, T., Moller, T., and Hagebeuker, B. Multidimensional Measurement by Using 3-D PMD sensors. *Advances in Radio Science*, 5 (2007), 135-146.